brium. This spontaneous reversibility provides further evidence against the suggestion (Kaempfer, 1973) that the free 70S ribosome is a storage form, appearing only under conditions of impaired protein synthesis.

References

Azzam, M. E., Perazzolo, C. A., and Algranati, I. D. (1972), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 21*, 165.

Beller, R. J., and Davis, B. D. (1971), *J. Mol. Biol. 55*, 477.

Davis, B. D. (1971), *Nature (London), 231*, 153.

Davis, B. D., and Mingioli, E. S. (1950), *J. Bacteriol. 60*, 17.

Dubnoff, J. S., and Maitra, U. E. (1971), *Methods Enzymol. 20*, 248.

Fiske, C. H., and Subbarow, Y. (1925), *J. Biol. Chem. 66*, 375.

Infante, A. A., and Baierlein, R. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 1780.

Iwasaki, K., Sabol, S., Wahba, A. J., and Ochoa, S. (1968), *Arch. Biochem. Biophys. 125*, 542.

Kaempfer, R. (1970), *Nature (London) 228*, 534.

Kaempfer, R. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 2458.

Kaempfer, R. (1973), *J. Mol. Biol. 71*, 583.

Parenti-Rosina, R., Eisenstadt, A., and Eisenstadt, J. M. (1969), *Nature (London) 221*, 363.

Sabol, S., Meier, D., and Ochoa, S. (1973), *Eur. J. Biochem. 33*, 332.

Sabol, S., and Ochoa, S. (1971), *Nature (London), New Biol. 234*, 233.

Sabol, S., Sillero, M. A. G., Iwasaki, K., and Ochoa, S. (1970), *Nature (London) 228*, 1269.

Subramanian, A. R. (1972), *Biochemistry 11*, 2710.

Subramanian, A. R., and Davis, B. D. (1970), *Nature (London) 228*, 1273.

Subramanian, A. R., and Davis, B. D. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 2453.

Subramanian, A. R., and Davis, B. D. (1973), *J. Mol. Biol. 74*, 45.

Subramanian, A. R., Davis, B. D., and Beller, R. J. (1969), *Cold Spring Harbor Symp. Quant. Biol. 34*, 224.

Subramanian, A. R., Ron, E. Z., and Davis, B. D. (1968), *Proc. Nat. Acad. Sci. U. S. 61*, 761.

Zitomer, R. S., and Flaks, J. G. (1972), *J. Mol. Biol. 71*, 263.

Collapsed Structure Polymers. A Scattergun Approach to Amino Acid Copolymers†

S. Prabhakara Rao, David E. Carlstrom, and Wilmer G. Miller*

ABSTRACT: A scheme is proposed for surveying the requirements of producing nonnatural sequence, globular, space-filling polyamino acids based on copolymerization and a double chromatographic procedure. Both experimental and theoretical considerations are used to show that the chromatographic technique can distinguish material having significant tertiary structure from material having only secondary structure. However, in a conformationally heterogeneous mixture the extent of tertiary structure and solvent exclusion is not readily deduced without auxillary measurements. A brief survey of several preparations is reported. A terpolymer containing equimolar amounts of lysine and glutamic acid plus 40 mol % alanine gives good evidence of having significant amounts of collapsed or partially collapsed material at neutral pH. Finally, arguments are made which suggest that the sequence requirements for globular formation deduced from studies on contemporary proteins may be overly strict.

Globular proteins and enzymes in biologically active conformations are unique macromolecular copolymers in that they apparently exist as collapsed, space-filling structures while molecularly dispersed in solution. Studies on wet crystals of globular proteins by X-ray diffraction indicate that they have a well-defined though irregular surface in contact with solvent (Blake, 1972; Drenth *et al.*, 1971; Quiocho and Lipscomb, 1971). The nonpolar amino acid side chains tend to be partially or completely removed from contact with solvent. Internal solvent when present is localized and not like bulk solvent. When molecularly dispersed in solution, hydrodynamic (Tanford, 1961; Yang, 1961) and chromatographic (Ackers, 1970) studies indicate the globular structure is maintained, while many types of studies indicate that some residues are not in contact with solvent (Roberts and Jardetzky, 1970; Tanford, 1962a, 1968; Wetlaufer, 1962). Any agent which destroys the globular integrity of a protein either destroys or greatly reduces the biological activity. As biological activity is sometimes lost with no apparent change in gross molecular conformation, loss of globular integrity is more than a sufficient condition for loss of activity in contemporary proteins and enzymes. By contrast, most synthetic polymers (Flory, 1953) as well as "completely denatured" proteins (Lapanje and Tanford, 1967; Tanford *et al.*, 1966, 1967) exist in solution as polymer random coils, in which most of the polymer domain is occupied by mobile, bulklike solvent and the entire monomeric unit has contact with solvent. The remaining type of macromolecular conformation in solution, an ordered structure such as a helix or a pleated sheet, has solvent in contact with at least part of each monomeric unit. The exclusion of solvent while molecularly dispersed is thus unique to globular proteins. Although there is considerable evidence that the globular or space-filling conformation is an important aspect of the biological, particularly catalytic, activity of proteins, this is not to say that a globular conformation is a static struc-

† From the Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455. *Received September 26, 1973.*

TABLE I: Some Approaches for Investigation of Globular Polypeptide Formation.

Denaturation–renaturation of globular proteins
Selective cleavage of globular proteins
Grafting (peptides or sugars) onto a globular protein
Specific sequence synthetic polypeptide
Multiple sequence synthetic polypeptides

ture. Side-chain mobility has been observed (Allerhand *et al.*, 1971), as well as large scale motions upon binding a substrate (Quiocho and Lipscomb, 1971). The evolution of proteins to their current state is a topic of much activity and discussion (Helsinki and Yanofsky, 1966; Ponnamperuma, 1972), but an understanding of the factors involved in a polypeptide chain assuming a globular conformation is an important problem in itself.

The theories or models which have been proposed for globular protein stability can be classified conveniently as zero through third order, depending on whether they depend only on composition (zero order), or take into account amino acid sequence or primary (first) as well as secondary (second) and tertiary (third) structure. Important concepts such as hydrophobic interactions and entropy considerations, as deduced from protein denaturation and model compound studies (Kauzmann, 1959; Tanford, 1968, 1970), or nonpolar inner, polar outer size and shape relationships (Fisher, 1964; Waugh, 1954) are in practice zero order. Most, if not all, of what is known about globular stability comes from studies on proteins known to be globular, and hence rationalizes rather than predicts globularity. An important question becomes "What is the lowest level model needed to predict globular stability?" Zero-order theories such as side-chain contributions to hydrophobic interactions can be used to rationalize the globular stability of myoglobin and β-lactoglobulin (Tanford, 1962b), for example, but the random coil nature of β-caseins (Tanford, 1970) is not explained by the same considerations. The considerable experimental evidence on the reversibility of the denaturation of a number of proteins suggests that the globular conformation of contemporary proteins is determined by the primary structure (Helinski and Yanofsky, 1966; Tanford, 1968). Mutations leading to a single amino acid replacement can result in an inactive enzyme, although the large interspecies variation in amino acid sequences in polypeptide chains carrying out the same biological function suggests that many regions of a protein have little effect on biological activity or globularity.

The similarity of the backbone conformation of polypeptide chains carrying out the same biological function (Dayhoff, 1972), but from varying species, suggests that mutations leading to significant alteration of the premutant backbone conformation are not viable. This may involve a loss of biological activity, and hence be lethal, even though the globular structure may remain. Removal of small fragments from the C-terminal end of ribonuclease, for example, leaves the molecule globular although the catalytic activity may be altered (Anfinsen, 1956; Puett, 1972). Folding up into a globular structure appears to be a necessary though not sufficient condition for efficient catalysis by proteins. The amino acid sequences found in contemporary proteins may in fact be a very small fraction of the sequences which can fold up even under physiological conditions. With 18 amino acids there are 18^N possible sequences of N residues. Since contemporary organisms are thought to contain only 10^3–10^6 different proteins (Dayhoff, 1972), the total number of sequences currently being produced presents a negligible fraction of the total possible sequences.

Although a globular conformation may be completely determined by the amino acid sequence, it seems clear that any theory with predictive ability must take into account short- and long-range interactions, and hence must be a higher order theory. Recent work indicates that the secondary structure of proteins known to be globular may be rationalized on the basis of primary structure (Gusso, 1965; Lewis *et al.*, 1970, 1971; Lim, 1972; Ptitsyn and Finkelstein, 1970). The instability and decrease in secondary structure resulting from removal of short segments (Atassi and Singhal, 1970; Singhal and Atassi, 1970) suggest that third-order theories are necessary to predict the stability of arbitrary sequences. Attempts along this line using potential functions seem doomed to failure for high molecular weight polypeptides due to the large number of possible conformations and the precision in potential energy functions needed to find a minimum energy structure. Additionally, a globular protein may, in fact, not be a minimum energy structure.

It seemed to us that an experimental approach might be useful. In this article we explore approaches to the investigation of collapsed, globular polypeptides, and report on some experimental attempts to produce synthetic polypeptides with collapsed structure.

Investigative Approaches

There are several experimental approaches one can suggest for exploring the requirements necessary to produce molecularly dispersed, globular proteins. Some of these are shown in Table I. The first four use contemporary globular proteins as the starting point. Denaturation studies currently contribute significantly, but incompletely, to our understanding of globular structures. The second approach, selective cleavage, compares a slightly altered protein (Anfinsen, 1956; Puett, 1972) or protein fragment (Atassi and Singhal, 1970; Singhal and Atassi, 1970) with the unaltered protein. In addition to showing the subtle nature of long-range interactions in stabilizing the globular structure, this approach yields important evidence that the folding up into a globular structure is not biologically controlled. The third approach, grafting of peptides or sugars onto globular proteins, explores the effect of additions on protein conformation. Most efforts of this nature have been oriented toward understanding enzymic activity rather than adding quantitiatively to an understanding of globular stability (Goldstein, 1972; Katchatski *et al.*, 1964). The fourth approach, a specific sequence synthesis such as the Merrifield solid state synthesis (Merrifield, 1969), seems especially useful for exploring the effect of altering particular chain positions in an amino acid sequence known to fold up into a globular structure. Due to the multiplicity of possible sequences it would not seem a particularly rewarding approach to a general exploration of the requirements for globular behavior.

In each of the first four methods one uses a known globular structure, *i.e.*, a native protein, as a reference point. This, of course, restricts one to contemporary protein sequences. The fifth approach, multiple sequence synthesis or copolymerization, might be called the "scattergun" approach in contrast to the "single shot" approach to the Merrifield synthesis. It has the advantage that many sequences can be explored simultaneously. It has the disadvantage that there are no reference

points, and that the distribution of sequences may be so large that each molecule in the system has a high probability of being unique. With respect to the latter point this type of synthesis is commonly used by those interested in prebiotic proteins, and the amino acid distribution in these preparations has been shown to be nonrandom (Fox, 1965, 1969; Helinski and Yanofsky, 1966). This will reduce the number of sequences (*cf.* sequences). The question of a reference point is of major importance. How does one decide whether or not a mixture of polypeptides contains a significant number of molecules with collapsed or substantially collapsed structure? After considering several approaches we decided that a double chromatographic procedure offered the most potential.

Analytical gel permeation chromatography (GPC)[1] of globular proteins is a thoroughly investigated and much used technique (Ackers, 1970; Determann, 1967). If all globular proteins were spherical and of constant density, there would be a linear relationship between the elution volume and the logarithm of the molecular weight. Figure 1, curve 1, is the relationship between molecular weight and spherical radius assuming a partial specific volume of 0.73 $cm^3\ g^{-1}$. Globular proteins are hydrated and also deviate from spherical symmetry and in practice the semilogarithmic relationship is observed only over a limited range. There appears to be a better correlation between the elution volume and the equivalent hydrodynamic (Stokes) radius (Ackers, 1970). Shown in Figure 1 is the molecular weight dependence of the Stokes radius for a number of globular proteins (filled circles), deduced from diffusion coefficient measurements (Ackers, 1964; Laurent and Killander, 1964), as well as radii deduced from X-ray measurements (×) taken as $(5/3)^{1/2}$ times the radius of gyration (Tanford, 1961). The highest molecular weight value is that for bushy stunt virus and is included here with the globular proteins because of its high molecular weight, spherical nature, and partial specific volume close to that of proteins. Curve 3, Figure 1, is the extrapolation (r^3_{Stokes} = 0.64 MÅ) of the experimentally determined Stokes radii. Giddings *et al.* (1968) in their theoretical study of the GPC of rigid particles concluded that the "mean external length," $\bar{L}$, was a better parameter than molecular weight, equivalent hydrodynamic radius, or radius of gyration. They predict that two rigid particles with the same $\bar{L}$ have the same elution volume, irrespective of shape. Globular proteins sometimes deviate significantly from spherical shape and are usually described as ellipsoids. For the purpose of mathematical simplicity we will consider capsule shaped molecules rather than ellipsoids. The mean external length of a capsule shaped molecule (cylinder with hemispherical ends) is related to that of a spherical shaped molecule of the same density by the simple relationship

$$\bar{L}_{capsule} = [(1 + x)/(2 + 6x)^{1/3}]\ \bar{L}_{sphere} \qquad (1)$$

where x is the ratio of length to diameter of the capsule and $\bar{v}_2$ = constant. Taking $x = 2$ as a reasonable asymmetry results in

$$\bar{L}_{capsule} = 1.24\ \bar{L}_{sphere} \qquad (2)$$

Curve 2 in Figure 1 was generated from eq 2 with $\bar{v}_2 = 0.73$ as in curve 1. Curves 1–3 differ, but not greatly. There is little doubt that a column calibrated with globular proteins of known Stokes radii and molecular weights will give meaningful rigid particle sizes and approximate molecular weights,

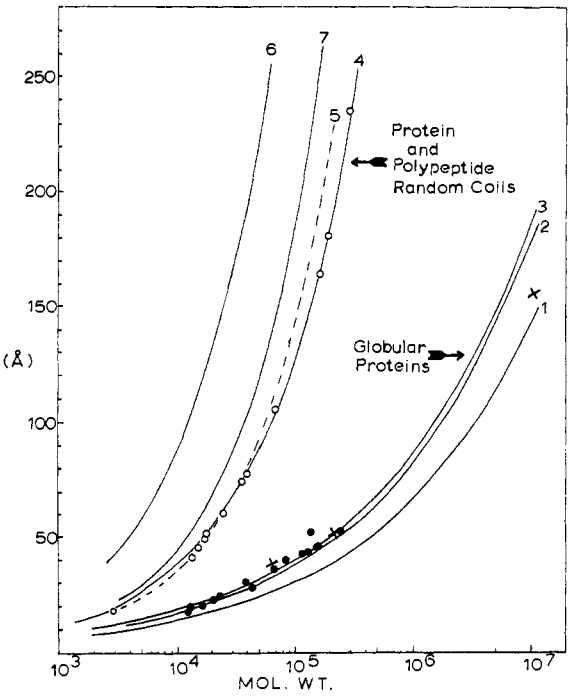

FIGURE 1: Molecular size (ångströms) as a function of molecular weight: (1) radius for spherical molecule with $\bar{v}_2 = 0.73$; (2) mean external length of capsule-shaped molecule, $x = 2$, $\bar{v}_2 = 0.73$ (eq 2); (3) equivalent sphere Stokes radii (r^3_{Stokes} = 0.64 mol wt Å) of globular proteins extrapolated from experimental values [(●) Laurent and Killander, 1964; Ackers, 1964; (×) Tanford, 1961, p 307]; (4) root-mean-square radius of gyration (R_G) of protein random coils in 6 M Gdn·HCl (experimental values: (○) Tanford *et al.*, 1966; Reisner and Rowe, 1969); (5) R_G of polyglutamic acid, pH 7, 0.4 ionic strength; (6) $\sqrt{\langle r^2 \rangle}$ of polyglutamic acid, pH 7, 0.4 ionic strength (Hawkins and Holtzer, 1972); (7) R_G of polypeptide random coil with a residue weight of 105 and an expansion coefficient appropriate to polyglutamic acid at pH 7, 0.1 ionic strength.

whose further interpretation depends on what is known or assumed about the geometric asymmetry (*cf.* sequences).

At the opposite end of the conformational hierarchy is the polymer random coil, the most typical conformational state of a synthetic polymer. Much is known about the mean conformation of both synthetic polypeptide random coils (Brant and Flory, 1965a,b; Hawkins and Holtzer, 1972; Miller *et al.*, 1967) and protein random coils (Lapanje and Tanford, 1967; Miller and Goebel, 1968; Reisner and Rowe, 1969; Tanford *et al.*, 1966, 1967). Studies on linear, branched, comb, and star homopolymers and block and graft copolymers show (Benoit *et al.*, 1966; Grubisic *et al.*, 1967b) convincingly that the elution of random coil polymers on GPC columns is related to the hydrodynamic volume of the molecule, *i.e.*, it is related to $[\eta]M$. Theoretical considerations support these experimental observations (Casassa and Tagami, 1969). For linear random coil polymers eq 3 is applicable (Flory, 1953), where Φ is a constant and $\langle r^2 \rangle$ is the mean-square end-to-end distance. Inasmuch as $\langle r^2 \rangle^{3/2}$ has the dimensions of volume, $[\eta]M$ is a measure of the volume of the polymer domain. The volume of a random coil is a nebulous quantity and it becomes crucial to know what aspect of a random coil's "volume" is important for GPC. This can be deduced experimentally by taking a column calibrated with globular proteins of known Stokes radii and chromatographing polypeptide random coils. Taking the well-characterized random coil sodium polyglutamate

$$[\eta] = \Phi\langle r^2 \rangle^{3/2}/M \qquad (3)$$

[1] Abbreviations used are: GPC, gel permeation chromatography; Gdn·HCl, guanidine hydrochloride.

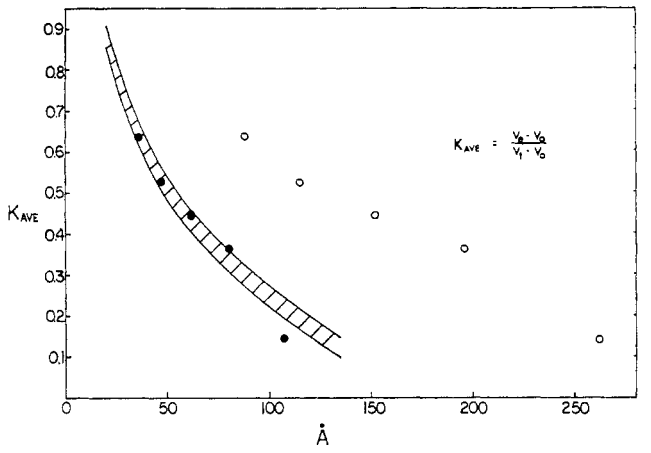

FIGURE 2: Comparison of elution position (K_{AVE}) for globular proteins as a function of equivalent Stokes radius (▨) with K_{AVE} for polyglutamic acid random coils as a function of R_G (●) and of $\sqrt{\langle r^2 \rangle}$ (○).

at pH 7 in 0.4 M NaCl (Hawkins and Holtzer, 1972) the root-mean-square end-to-end distances were calculated. The values, smoothed and extrapolated into the low molecular weight region, are given in Figure 1, curve 6. What would seem to be an intuitively more meaningful measure of volume, the radius of gyration, R_G, is shown in curve 5 computed through the relationship $\langle r^2 \rangle$ equals $6R_G^2$.

Let us now consider Bio-Gel A-5m, an agarose gel stated by the manufacturer (Bio-Rad Laboratories) to exclude globular proteins greater than 5×10^6 in molecular weight and to fractionate globular proteins down to a mol wt of 10,000. If we use the extrapolated Stokes radii, curve 3, as an indicator of pore size, we would predict that sodium polyglutamate random coils greater than about 25,000 would be excluded if $\langle r^2 \rangle^{1/2}$ was the appropriate parameter, whereas the exclusion limit would be about 100,000 if R_G was the appropriate parameter. Previously fractionated and characterized polyglutamates (Snipp, 1964; Snipp *et al.*, 1965) were passed through a Bio-Gel A-5m column in 0.5 M NaCl (pH ~7). A sample of mol wt 107,000 was excluded whereas a sample of 65,000 was accepted. A plot of K_{AVE} as a function of R_G is nearly superimposable with a plot of K_{AVE} as a function of the Stokes radii of globular proteins, Figure 2, quite in contrast to the behavior of $\langle r^2 \rangle^{1/2}$.

These results are in accord with the studies of Fish *et al.* (1969) on protein random coils in 6 M guanidine hydrochloride (Gdn·HCl) on Bio-Gel A-5m. They found a useful fractionation range of 80,000–1000, although the functional form of the molecular size dependence of the elution volume appears to be somewhat altered by the presence of 6 M Gdn·HCl (Ackers, 1970; Fish *et al.*, 1969). Shown in Figure 1 are the protein random coil radii of gyration (open circles) deduced from viscosity measurements (Lapanje and Tanford, 1967; Reisner and Rowe, 1969; Tanford *et al.*, 1966, 1967) in 6 M Gdn·HCl and curve 4 is the smoothed curve through these data. Using curves 3 and 4 one predicts an exclusion limit of slightly over 100,000 and a lower fractionation limit of about 2000. Again the use of Stokes radii for globular structures and R_G for random coil polymers seems meaningful.

Let us now take a Sephadex G-75 column (Pharmacia Fine Chemicals) with solvent conditions such that contemporary proteins are globular, place on it a mixture of globular and random coil polypeptides, and collect fractions at intervals of equal Stokes radii as shown in Figure 3A. If each of these fractions were now put onto a Bio-Gel A-5m column with 6 M

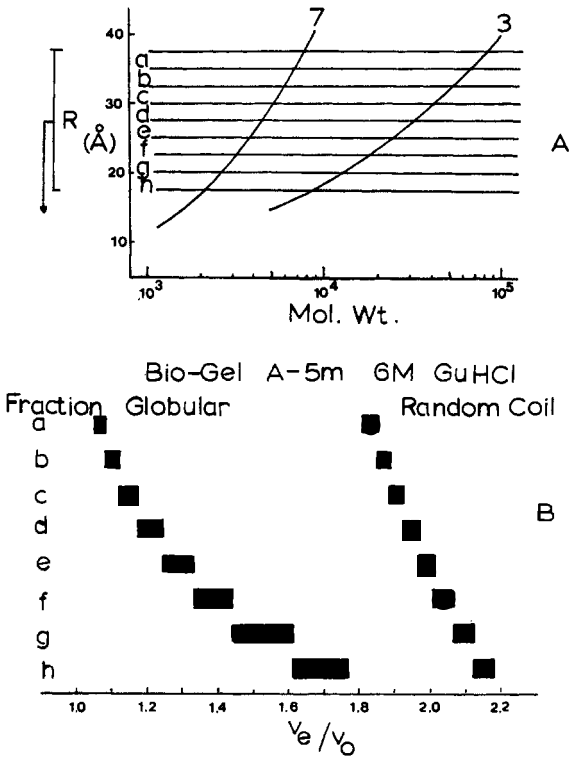

FIGURE 3: Schematic elution (B) of fractions from a globular column when placed on a column containing a denaturing solvent. Curves 3 and 7 in A are portions of 3 and 7 in Figure 1.

Gdn·HCl as a medium, the elution positions would be as indicated in Figure 3B. There is no doubt that this double chromatographic procedure can separate and distinguish globular from random coil polypeptides. Although there may be an entire spectrum of conformations present, we will show later that unless a polypeptide chain has substantial tertiary structure its elution position will look much more like a random coil than a globular structure. By making use of the available information on the important factors necessary to make a contemporary protein globular and a double chromatographic procedure one can synthesize and survey a wide variety of preparations with a minimum of effort. By the double chromatographic procedure one can isolate for further study material which appears to be globular under solvent conditions where contemporary proteins are globular.

Experimental Section

Polypeptide Synthesis. *N*-Carboxyanhydrides of amino acids, side chain blocked where necessary with benzyl or carbobenzoxy groups, were either purchased (Pilot Chemicals, Inc., or Miles Laboratories) or were prepared from the amino acid and phosgene by standard procedures (Farthing, 1950). The anhydrides, in the desired mole ratios, were dissolved in a 1:4 (v/v) mixture of dry dioxane and benzene. The polymerization was initiated with sodium methoxide, diisopropylamine, or triethylamine. The reaction was generally monitored by CO_2 evolution. The resulting polymers were deblocked with either dry HBr or phosphonium iodide and hydrogen.

Column G-75. A K 15/90 column (Pharmacia) was packed with Sephadex G-75 (Pharmacia) using standard procedures (Determann and Michel, 1966). Sephadex G-75 was chosen as it fractionates globular proteins in the mol wt range 10,000–80,000. The eluting solvent was 0.025 M phosphate buffer at

pH 7, unless stated otherwise. The flow rate was regulated by a pump (Technicon no. 1) at a nominal flow rate of 13 cm^3 hr^{-1}. The column effluent was passed through a variable path length flow cell (Beckman) in a Cary 15 spectrophotometer and monitored at 210 nm, except when monitoring blue dextran (630 nm) or 2,4-dinitrophenylalanine (360 nm). The column was calibrated by measuring the elution volume of standard globular proteins (Pharmacia). The void volume (V_0) was determined with blue dextran 2000 (Pharmacia) and the inner volume (V_i) with 2,4-dinitrophenylalanine. In a typical experiment, with both standards and samples, 1–2 mg of material dissolved in 0.5 ml of buffer was injected onto the column.

Column A-5m. A SR 25/45 column (Pharmacia) was packed with Bio-Gel A-5m, 100–200 mesh (Bio-Rad Laboratories, lot 74423), which had been preconditioned in 6 M Gdn·HCl. A hydrostatic pressure of less than 10 cm of solvent was used when packing as well as during column operation. The eluting solvent was 6 M Gdn·HCl with flow regulation as with column G-75. The column was calibrated as with column G-75 except that the proteins had been converted to random coils by incubation with either 0.05 or 0.1 M mercaptoethanol in 6 M Gdn·HCl (pH 8.4) prior to placement on the column. The random coil proteins were injected onto the column and eluted with 6 M Gdn·HCl containing the same mercaptoethanol concentration as was used to reduce the disulfide bonds. The column effluent was monitored spectrophotometrically (280 nm) or turbidimetrically depending on whether the reducing agent was 0.05 or 0.1 M, respectively. The elution positions were found to be independent of the mercaptoethanol concentration, thus making the spectrophotometric procedure practical. Since none of the synthetic polypeptides to be reported on here contained cysteine, mercaptoethanol was omitted when the synthesized material was chromatographed, and the effluent was then monitored at 220 nm. Elution positions were determined volumetrically and not gravimetrically. Elution volumes were in good agreement with those of Fish *et al.* (1969). As with column G-75, 1–2 mg of sample in 0.5 ml of solvent was the typical charge.

Other Measurements. Circular dichroism measurements were made using a Cary 60 spectropolarimeter. A Corning Model 112 digital pH meter was employed for pH determinations. Viscosities were determined with Cannon-Ubbelohde semimicro dilution viscometers.

Experimental Results

Homopolymers. In considering the use of GPC, we discussed earlier the elution of random coil polyglutamate in 0.5 M NaCl on Bio-Gel A-5m, Figure 2. We concluded that the radius of gyration was the appropriate random coil parameter. Using curves 3 and 5, Figure 1, the exclusion limit for random coils in 0.5 M NaCl on G-75 would be about 10,000, the lower fractionation limit for globular proteins on this column. In 0.025 M phosphate the chain will be somewhat more expanded due to electrostatic interactions. Curve 7, Figure 1, is the radii of gyration appropriate to the polymer at 0.1 ionic strength (Hawkins and Holtzer, 1972), but calculated on the basis of a residue weight of 105. Conversion to the residue weight of glutamate will lower the R_G values, though the slightly lower ionic strength employed in column G-75 will counter this. Using curves 3 and 7 we would predict an exclusion limit of about 8000 for random coils on G-75. Shown in Figure 4 is the elution profile on column G-75 of a polyglutamate sample with a maximum weight average molecular

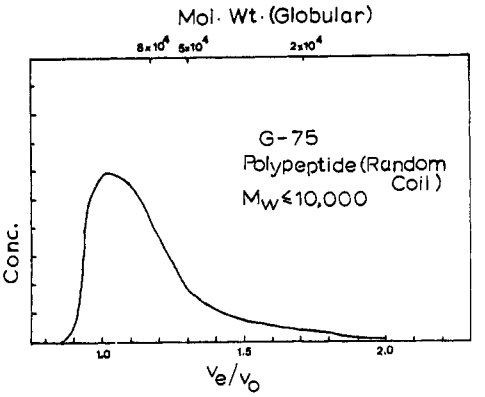

FIGURE 4: Chromatogram of polyglutamic acid (pH 7), in 0.025 M phosphate on Sephadex G-75.

weight of 10,000 (Rao and Miller, 1973), which had not been previously fractionated. Much of the sample is excluded, or nearly so.

Binary Copolymers. Copolymers containing a hydrophobic side chain such as leucine or alanine, and an ionic side chain such as lysine or glutamate, were not soluble at pH 7 when the mole fraction of hydrophobe was large. At low percentage hydrophobe these polymers are known to be random coil (Fasman *et al.*, 1964; Miller and Nylund, 1965; Nitta and Sugai, 1967; Sugiyama and Noda, 1970) and seemed uninteresting.

Copolymers of lysine and glutamate ranging from 25 to 75% lysine are quite insoluble at pH 7 and over a considerable range of pH on either side of neutrality. At pH 11 or above they were soluble but excluded from G-75. Further studies were not made.

Copoly(Asp,Glu,Lys,Ala,Leu,Trp,Tyr,Val). *N*-Carboxyanhydrides were copolymerized using the monomer composition 10, 27, 31, 6, 4, 8, 10, and 4 for Asp, Glu, Lys, Ala, Leu, Trp, Tyr, and Val, respectively. The resulting copolymer, after side-chain deblocking with phosphonium iodide and hydrogen, was dissolved in a 5% $NaHCO_3$ solution, dialyzed, and lyophilized. Amino acid analysis gave a composition of 10.5, 30.5, 46.2, 5.5, 4.3, 0, 0, and 2.9. Tyrosine and tryptophan were not incorporated. This copolymer was insoluble in the pH range 4–11. At pH values where the polymer was soluble, it was almost totally excluded from G-75.

Copoly(Asp,Glu,Lys,Orn,Gly,Leu,Trp,Tyr,Val). *N*-Carboxyanhydrides were copolymerized using the monomer composition 8, 22, 26, 3, 8, 3, 9, 9, and 12 for Asp, Glu, Lys, Orn, Gly, Leu, Trp, Tyr, and Val, respectively. The resulting copolymer was worked up as with the preceding polymer. Amino acid analysis gave a composition of 9.5, 25.6, 46.5, 0, 7.9, 3.6, 0, 0, and 7.0. Tyrosine and tryptophan were again not incorporated, nor was ornithine. The pH solubility and behavior on G-75 was analogous to that found with the preceding copolymer.

Copoly(Glu,Lys,Ala). *N*-Carboxyanhydrides were copolymerized with diisopropylamine as initiator using the monomer composition 30, 30, and 40 for Glu, Lys, and Ala, respectively. After deblocking and work-up as with the preceding polymers, amino acid analysis indicated a composition of 30.0, 31.1, and 38.9. This polymer was quite soluble at all pH values.

The copolymer was accepted onto the G-75 column at pH 7, Figure 5A, as well as on the A-5m column in 6 M Gdn·HCl, Figure 6. The effluent from the G-75 column was collected into two fractions, the cut being made at a volume corresponding to $V_e/V_0 = 1.4$. The two fractions were isolated

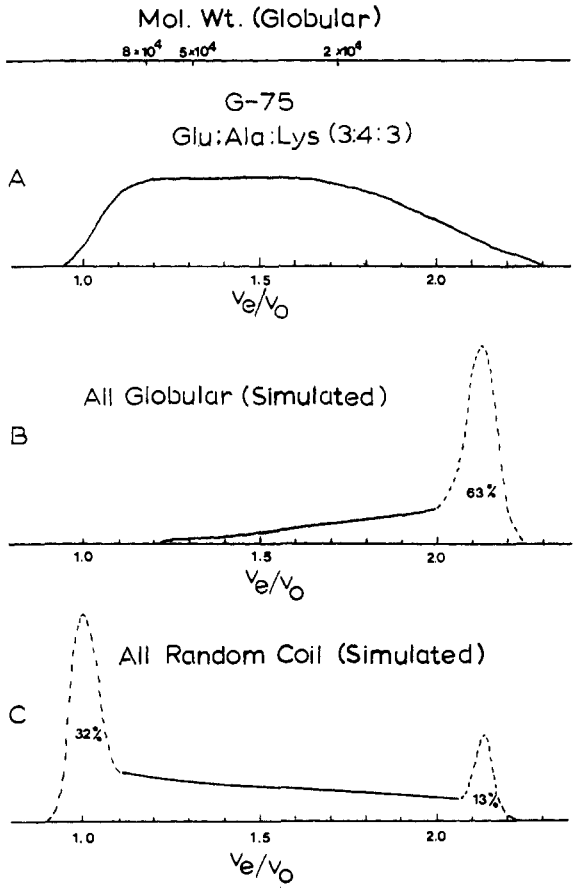

FIGURE 5: Chromatogram of copoly(Glu,Lys,Ala) (3:3:4) on Sephadex G-75, pH 7, in 0.025 M phosphate (A). Simulated chromatogram if terpolymer were completely globular (B) or completely random coil (C).

and separately placed on the A-5m column. The first fraction was eluted earlier than the unfractionated material, with a maximum at $V_e/V_0 = 1.75$. The second fraction shifted to high elution volumes with a maximum at $V_e/V_0 = 2.12$. The copolymer was also chromatographed at pH 11 on the G-75 column. It was now predominantly excluded from the gel and was similar in appearance to Figure 4.

In 6 M Gdn·HCl the terpolymer had an intrinsic viscosity of 17.0 ± 0.5 $cm^3\ g^{-1}$, while in 0.025 M phosphate (pH 7) the intrinsic viscosity was 12.6 ± 0.1 $cm^3\ g^{-1}$.

Circular dichroism spectra of the unfractionated terpolymer as well as of the two fractions at pH 7 in 0.025 M phosphate show the double minimum characteristic of the α helix. The helical content, estimated from the ellipticity at 221 nm, was 46% for the unfractionated material and for fraction two, and slightly higher for fraction one.

Effect of Secondary Structure on Chain Dimensions and Gel Filtration

In the absence of any ordered secondary structure, *i.e.*, for the polymer random coil, the experimental chain dimensions of synthetic polypeptides and protein random coils may be rationalized easily using the isomeric rotation state approach (Brant and Flory, 1965b; Miller *et al.*, 1967; Miller and Goebel, 1968). If the molecules become completely ordered intramolecularly so as to have an intact, α-helical conformation the root-mean-square end-to-end distance ($\sqrt{\langle r^2 \rangle}$) is the same as the length L of the molecule and R_G becomes $L/\sqrt{12}$. The mean external length $\bar{L}$ (Giddings *et al.*, 1968) is $0.5(L + D)$ for long, thin rods, where D is the helix diameter. $\bar{L}$ may be computed alternatively by eq 2 or for the helix taken as an ellipsoid of revolution. Experimentally helices have been shown to chromatograph according to their hydrodynamic volume (Grubisic *et al.*, 1967a). Irrespective of whether $\bar{L}$ or R_G is the appropriate size parameter, both are greater than R_G for the corresponding molecular weight random coil so long as the molecular weight is above a few thousand. Thus, any fully helical molecules, if present, will not be confused with globular material in the double chromatographic procedure.

Partial rather than full helicity is a much more probable secondary structure. The chain dimensions (R_G or $\sqrt{\langle r^2 \rangle}$) of homopolymers containing partially helical secondary structure, but no tertiary structure, may be computed assuming Ising model statistics to describe the conformational population of helix and random coil states (Miller and Flory, 1966). Viscosity measurements (Hayashi *et al.*, 1969) through the helix–coil transition are in qualitative agreement with these computations. The validity of the calculated dimensions is most impressively demonstrated by the agreement between the Ising model cooperativity parameter determined from hydrodynamic and from optical measurements (Hagnauer and Miller, 1970; Bychkova *et al.*, 1971).

The calculation of chain dimensions may be extended easily to partially helical copolymers if an appropriate description of the helix–random coil transition is available. Assuming the Zimm–Bragg (1959) or equivalent (Zimm and Rice, 1960) description of the helix–coil transition to be applicable the configurational partition function Z is given for a chain of n peptide bonds by

$$Z = [1\ 0]U_1U_2\cdots U_i\cdots U_n\begin{vmatrix}1\\1\end{vmatrix} \quad (4)$$

where we have assumed a 2 × 2 statistical weight matrix of the form

$$U_i = \begin{vmatrix}1 & \sigma s\\ 1 & s\end{vmatrix}_i \quad (5)$$

The characteristic ratio, $\langle r^2 \rangle / nl^2$, where l is the virtual bond length (3.8 Å), is given (Miller and Flory, 1966; Flory, 1969) by

$$\langle r^2 \rangle / nl^2 = 1 + (2/Znl^2)(1\ 0\cdots 0)G_1G_2\cdots G_i\cdots G_{n-1}[0\ 0\ 3.8\ 0\ 0\ 3.8\ 0\ 0\ 1\ 1]^T \quad (6)$$

where

$$G_i = \begin{vmatrix}U & (U \otimes l^T)||T|| & 0\\ 0 & (U \otimes E_3)||T|| & U \otimes l\\ 0 & 0 & U\end{vmatrix}_i \quad (7)$$

In eq 6 and 7 0 is the null matrix, E_3 is the identity matrix of third order, $\otimes$ is the direct matrix product, l^T is the row matrix [3.8 0 0], l is the transpose of l^T, and $||T||$ is the pseudo-diagonal matrix given in eq 8, whose "elements" are 3 × 3

$$||T|| = \begin{vmatrix}\langle T_c \rangle & 0\\ 0 & T_h\end{vmatrix} \quad (8)$$

coordinate transformation matrices for averaged random coil ($\langle T_c \rangle$) or helical (T_h) residues. In the case of a terpolymer there are three U_i and G_i matrices, one for each type of residue. The mean-square radius of gyration may be computed directly, or can be obtained more easily from the relationship $R_G^2 = \langle r^2 \rangle / 6$, since previous calculations indicate that even

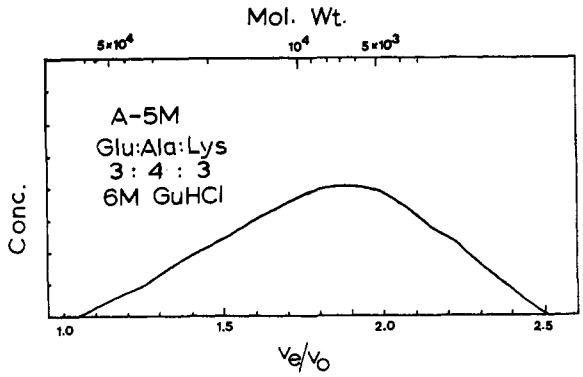

FIGURE 6: Chromatogram of copoly(Glu,Lys,Ala) (3:3:4) on Bio-Gel A-5m in 6 M Gdn·HCl.

a small percentage of random coil units leads to values of $\langle r^2\rangle/R_G{}^2$ very close to the statistical chain value of 6. The average fraction f_h of helical units in the chain is given by

$$f_h = (1/nZ)[1\ 0\ 0\ 0]\begin{vmatrix} U & U' \\ 0 & U \end{vmatrix}_1 \cdots \begin{vmatrix} U & U' \\ 0 & U \end{vmatrix}_i \cdots \begin{vmatrix} U & U' \\ 0 & U \end{vmatrix}_n \begin{vmatrix} 0 \\ 0 \\ 1 \\ 1 \end{vmatrix} \quad (9)$$

where

$$U_i' = \begin{vmatrix} 0 & \sigma s \\ 0 & s \end{vmatrix}_i$$

Other conformational averages, if desired, may be computed from equations similar to eq 9.

Any "randomly" polymerized copolymer contains many amino acid sequences. Inasmuch as the Glu:Lys:Als (30:30:40) terpolymer is the only polymer reported herein which gives evidence of collapsed structure, calculations were performed pertinent to this copolymer. The binary reactivity ratios indicated that the amino acid sequences were not random, *i.e.*, the binary reactivity ratios are not unity in our polymerization media (S. P. Rao, D. Carlstrom, and W. G. Miller, manuscript in preparation). From the binary reactivity ratios a stochastic matrix P was generated (Ham, 1964) (eq 10) whose elements are sequential probabilities appropriate to

$$P = \begin{array}{c} \\ \text{G} \\ \text{L} \\ \text{A} \end{array}\begin{array}{c} \begin{array}{ccc} \text{G} & \text{L} & \text{A} \end{array} \\ \begin{vmatrix} 0.30 & 0.31 & 0.39 \\ 0.31 & 0.25 & 0.44 \\ 0.13 & 0.16 & 0.71 \end{vmatrix} \end{array} \quad (10)$$

the Glu:Lys:Ala terpolymer with a 30:30:40 monomer feed. Thus, for example, the probability that a Glu is followed by a Glu, Lys, or Ala is 0.30, 0.31, or 0.39, respectively. Through the use of a random number generator, sequences may be generated in accord with the sequential probabilities given in eq 10. One can either do a Monte Carlo calculation, *i.e.*, for a given n, σ and s values generate chains, compute all desired properties, average, and keep adding chains until all properties fluctuate less than preassigned limits, or one can calculate the properties for each sequence generated and tabulate without averaging. In the first case the mean composition will be the same for each n, σ, and s and a smooth curve is calculated, whereas in the latter case the composition will vary. A σ value of 0.001 was assumed for each residue, and a chain length of 300. Shown in Figure 7 is the Monte Carlo line for s_{Ala} equal to 1.1 and variable s_{Glu} and s_{Lys} and a set of discrete points with s_{Ala} varying from 1.1 to 1.2, and s_{Glu} and s_{Lys} from 0.7 to 1.1. Several features emerge. Deviations from mean values are not large. Although the sequences are nonrandom, the mean helicity is more important in determining chain dimensions than the sequence, composition, or specific values assigned to the various s values, at least within the range of s values deemed reasonable. The variation of R_G with mean helicity is similar to the homopolymer results. Finally, we note that in the f_h range of interest, 0.3–0.6, the chain dimensions are never more than a few per cent different from the random coil values. The shallow minimum in R_G as helicity sets in becomes even more shallow as n becomes smaller. Thus, we feel confident in asserting that molecules with partially helical secondary structure will chromatograph very similarly to a random coil of the same molecular weight.

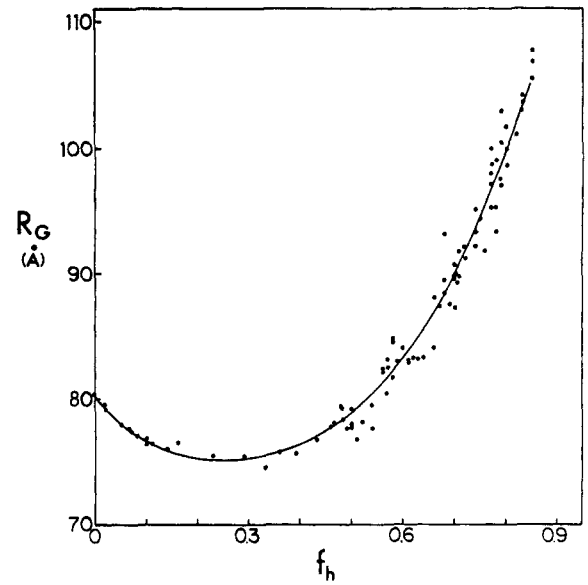

FIGURE 7: Effect of secondary structure (helical) on chain dimensions of terpolymer for a chain of 300 residues with σ = 0.001 and sequential probabilities given by eq 10: points, specific sequence chains (s_{Ala} = 1.1–1.2; s_{Lys} and s_{Glu} = 0.7–1.1); solid line, Monte Carlo calculation (s_{Ala} = 1.1; s_{Lys} and s_{Glu} = 0.5–1.2).

The intramolecular β structure is another type of secondary structure which must also be considered. If a chain bends and folds back on iself to give a short run of β structure, the chain dimensions will not be substantially different from those with partial helical structure, and to good approximation are equivalent to random coils with knots (Miller and Goebel, 1968). A β structure involving residues remotely located along the chain will more severely reduce the dimensions, and must be properly classified as a tertiary structure. The remoteness necessary before a β structure is classified as a tertiary rather than as a secondary structure must, of course, be arbitrarily defined. However, it need not concern us when considering the Glu:Lys:Ala copolymer (*cf.* sequences).

Discussion

The Scattergun Approach. The approach to investigation of collapsed structure polymers which we have espoused involves synthesis as well as a scheme for detection and isolation of material for further study. The fractionation of globular molecules on gel columns seems well understood. The data displayed in Figure 2 put the fractionation of random coils on a good foundation in terms of the random coil parameter which controls the elution position. It was necessary to establish this relationship since in the usual application of gel chromatography to random coil polymers calibration with or comparison to rigid, globular particles is not made. By the double chromatographic procedure molecules with tertiary structure should be distinguishable and isolatable from those with no

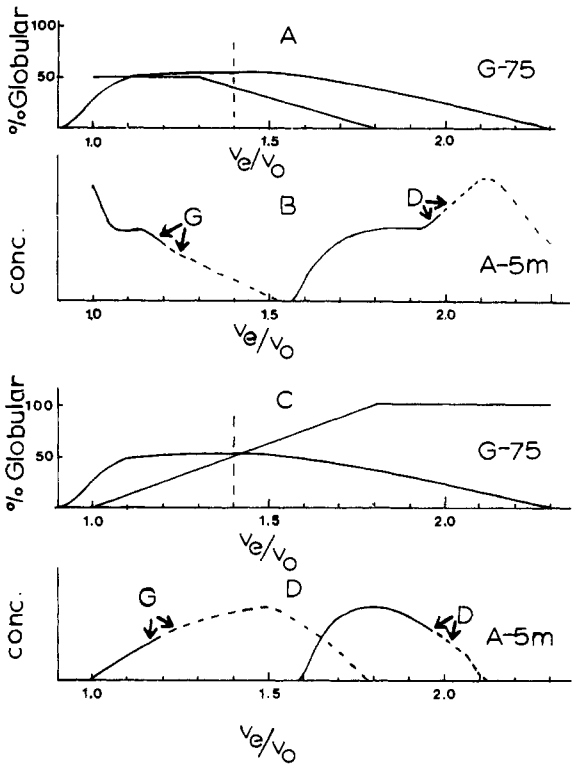

FIGURE 8: Simulated elution on Bio-Gel A-5m of terpolymer separated into two fractions on G-75 (fraction 1, $v_e/v_0 < 1.4$; fraction 2, $v_e/v_0 > 1.4$). Chromatogram B is simulated assuming the distribution of globular material on G-75 was as indicated by straight line segments in A; chromatograph D assumes distribution of globular material as indicated in C; solid lines in B and D are elution positions of fraction 1, and dashed lines of fraction 2; elution segments marked G (globular) and D (denatured) refer to conformation on G-75.

tertiary structure. It must be highly probable, however, that many molecules having collapsed structure are not in fact globular, relatively rigid, space-filling molecules. Further physical characterization must be made on the material showing collapsed structure in order to decide the degree of collapse and its rigidity. It is entirely possible that material showing a high degree of collapsed structure as its mean conformation is in fact widely fluctuating, and hence is not a tertiary structure in the current biochemical usage. This usage has, however, recently been challenged (Lakowicz and Weber, 1973).

Water-soluble globular proteins tend to have rather high contents of nonpolar as well as substantial amounts of positively and negatively charged side chains. These were our guiding synthetic principles for a brief survey. A major problem encountered was to obtain water-soluble polymers in the neutral pH range. This is not surprising as a high content of nonpolar groups undoubtedly will lead to aggregation and precipitation if the sequence is not correct for globular formation. It may even be a good screening criterion that a synthetic preparation containing a high nonpolar content and having protein-like solubility near its mean isoelectric point is indicative of globularity. The much enhanced solubility in the neutral pH range of equimolar glutamate-lysine terpolymers containing alanine in excess of their nonalanine counterparts thus seemed of interest. Indeed the 30:30:40 terpolymer was the only polymer in our survey which could be chromatographically investigated and which gave evidence of collapsed structure.

Glutamate-Lysine-Alanine Terpolymer. This copolymer has been studied previously (Morita *et al.*, 1967) but with respect to secondary structure only. They found 40–45% helicity at pH 8 and no evidence of β structure. Our measurement of secondary structure is in good agreement with their results.

Alanine is generally considered to be a marginal hydrophobe and certainly not as good as leucine or valine. Our finding of evidence for collapsed structure in a polymer containing 40% alanine as the only hydrophobe may then seem surprising. We would point out that poly(L-alanine) is extremely water insoluble, thus indicating that methyl–methyl contacts are highly favored compared to methyl–water contacts. Secondly, evidence has been presented (Ingwall *et al.*, 1968; Howard and Scheraga, 1972) which indicates that simple block copolymers of alanine and lysine have some tertiary structure. Tertiary structure in alanine copolymers is not then a precedent. What becomes of interest is the amount of tertiary structure, *i.e.*, the degree of collapse.

Some indication of the conformation may be obtained by taking the molecular weight distribution determined on the A-5m column, Figure 6, and simulating the results expected on the G-75 column. The simulated chromatograms, assuming completely globular or, conversely, material completely devoid of secondary and tertiary structure, are shown in Figure 5. Neither simulated chromatogram compares well with the observed one, Figure 5A. They do suggest that some of the high molecular weight material must have tertiary structure, and much of the low molecular weight material must be devoid of tertiary structure. Alternatively, a distribution of globular material may be assumed on the G-75 column and the expected chromatogram for the A-5m column predicted. Inasmuch as the elutant from the G-75 column was split into two portions and chromatographed on the A-5m column, giving chromatograms similar in shape to Figure 6 but with maxima shifted to higher and lower elution volumes, simulation of these can also be made. Results of such simulation are shown in Figure 8. The straight section lines in A and C are the assumed globularity distribution, and the corresponding simulated A-5m chromatographs are shown in B and D, respectively. The solid lines in B and D correspond to the elution positions from fraction 1, and the dashed lines correspond to positions from fractions 2. G refers to material which had been globular and D to material not globular in the G-75 column. Neither distribution of globular material gives simulated chromatographs in much agreement with that observed, which would indicate that the material with tertiary structure is only partially collapsed. Of course, one cannot take this kind of simulation too seriously as each molecular weight has been tacitly assumed to have a unique elution position, and not a band of finite width. The main conclusion to be drawn from this discussion is that the chromatographic procedure is a good technique for screening and isolation of material with tertiary structure, but that other methods must be used to determine the extent of collapse.

Prebiotic Proteins. Over the past 20 years much thought has been given to and experimental work carried out on the origin of amino acids and proteins, and of subsequent cellular development. Polyamino acids have been prepared under simulated prebiotic conditions by thermal polymerization (Fox, 1965, 1969; Saunders and Rohlfing, 1972), and also have been prepared by copolymerization of the *N*-carboxyanhydrides of the 18 commonly occurring amino acids (Hayakawa *et al.*, 1967). Polylysine obtained by thermal polymerization has been studied with respect to its catalytic ability and simulation of enzyme-like properties (Heinrich *et al.*, 1968, 1969). These

materials are often referred to as "protein-like" polymers and are frequently said to resemble contemporary proteins in many ways. But the one unique aspect of contemporary enzymes and some structural proteins, namely globularity, is to our knowledge never investigated in these studies. In the small amount of gel chromatography which has been done, there is little or no appreciation of the difference in size between globular and nonglobular material of the same molecular weight. The procedure we have outlined should make it easy to deduce if these heterogeneous preparations contain material with significant tertiary structure, and hence are indeed protein-like substances.

Probability of Success. In contrast to investigations on prebiotic polymers, those working with contemporary proteins are sharply focused on the requirement of globularity. It is frequently found that globularity and function are lost with the removal of only a few residues in a naturally occurring sequence. From analysis of the X-ray structure and amino acid sequence of proteins of similar function from varying species, some rather detailed ideas have been formulated with respect to viable mutations. These kinds of observations quite naturally lead to the idea of unique sequence requirements for globularity. From the point of view of contemporary protein research, the search for globular material in a randomly copolymerized sample has less chance of success than the proverbial search for the needle in a haystack. As we have stated earlier, we feel this may be too narrow a focus.

Theoretical considerations on the conformation and solution thermodynamics of synthetic polymers are of interest here. An approximate treatment of chain conformation into the region where long-range interactions are predominantly attractive predicts that in a poor solvent the polymer chain may precipitously exclude solvent from its domain and become globular (Ptitsyn *et al.*, 1968). However, these authors felt that this would lead to intermolecular interaction as well, and the polymer would likely precipitate before becoming globular. It has been suggested that it is problematical whether or not a chain of representative amino acids would offer even one globular conformation of low intramolecular energy (Flory, 1972). Theoretical considerations to the contrary, a paper has appeared recently which shows for the first time a synthetic homopolymer with unquestionable globular conformation (Ancifzieva *et al.*, 1972). The condition is simply a poor solvent. The polymer is not known to exhibit unusual properties in a good solvent. The predictions of Ptitsyn *et al.* (1968) are thus realized. Since there is nothing unique about this system and it is in accord with general predictions, it may be realized by a wide variety of homopolymers. The manner in which the chain is able to fold back on itself and form a space-filling structure is unknown, and seems to be currently beyond theoretical prediction. It does, however, suggest to us that the requirements for copolymers to form globular structures may not be as restrictive as is generally supposed.

References

Ackers, G. K. (1964), *Biochemistry 3*, 723.

Ackers, G. K. (1970), *Advan. Protein Chem. 24*, 343.

Allerhand, A., Doddrell, D., Glushko, V., Cochran, D. W., Wenkert, E., Lawson, P. J., and Gurd, F. R. N. (1971), *J. Amer. Chem. Soc. 93*, 544.

Ancifzieva, E. V., Volkenstein, M. V., Gottlieb, Y., Kzokovjak, M. G., Pacetov, W. D., Stepanov, B. B., and Skorkhodov, S. S. (1972), *Dokl. Akad. Nauk SSSR 207*, 224.

Anfinsen, C. B. (1956), *J. Biol. Chem. 221*, 405.

Atassi, M. Z., and Singhal, R. P. (1970), *J. Biol. Chem. 245*, 5122.

Benoit, H., Grubisic, Z., Rempp, P., Decker, D., and Zilliox, J. G. (1966), *J. Chim. Phys. Physiochim. Biol. 63*, 1507.

Blake, C. C. F. (1972), *Progr. Biophys. Mol. Biol. 25*, 85.

Brant, D. A., and Flory, P. J. (1965a), *J. Amer. Chem. Soc. 87*, 2788.

Brant, D. A., and Flory, P. J. (1965b), *J. Amer. Chem. Soc. 87*, 2791.

Bychkova, V. E., Ptitsyn, O. B., and Barkskaya, T. V. (1971), *Biopolymers 10*, 2161.

Casassa, E. F., and Tagami, Y. (1969), *Macromolecules 2*, 14.

Dayhoff, M. O. (1972), *in* Exobiology, Ponnamperuma, C., Ed., Amsterdam, North-Holland Publishing Co., pp 266ff.

Determann, H. (1967), Gelchomatographie, Berlin, Springer-Verlag.

Determann, H., and Michel, W. (1966), *J. Chromatogr. 25*, 303.

Drenth, J., Jansonius, J. N., Koekoek, R., and Wolthers, B. G. (1971), *Advan. Protein Chem. 25*, 79.

Farthing, A. C. (1950), *J. Chem. Soc.*, 3212.

Fasman, G. D., Lindblow, C., and Bodenheimer, E. (1964), *Biochemistry 3*, 155.

Fish, W. W., Mann, K. G., and Tanford, C. (1969), *J. Biol. Chem. 244*, 4989.

Fisher, H. F. (1964), *Proc. Nat. Acad. Sci. U. S. 51*, 1285.

Flory, P. J. (1953), Principles of Polymer Chemistry, Ithaca, N. Y., Cornell University Press.

Flory, P. J. (1969), Statistical Mechanics of Chain Molecules, New York, N. Y., Interscience.

Flory, P. J. (1972), Polymerization in Biological Systems, Ciba Foundation Symposium 7, Amsterdam, North-Holland Publishing Co., p 109.

Fox, S. (1965), *Nature (London) 205*, 328.

Fox, S. W. (1969), *Naturwissenschaften 56*, 1.

Giddings, J. C., Kucera, E., Russell, C. P., and Myers, M. N. (1968), *J. Phys. Chem. 72*, 4397.

Goldstein, L. (1972), *Biochemistry 11*, 4072.

Grubisic, Z., Reibel, L., and Spach, G. (1967a), *C. R. Acad. Sci., Ser. C 264*, 1690.

Grubisic, Z., Rempp, P., and Benoit, H. (1967b), *J. Polym. Sci. 135*, 753.

Gusso, A. V. (1965), *Biophys. J. 5*, 809.

Hagnauer, G., and Miller, W. G. (1970), *Biopolymers 9*, 589.

Ham, G. E. (1964), Copolymerization, Vol. 18, New York, N. Y., Interscience.

Hawkins, R. B., and Holtzer, A. (1972), *Macromolecules 5*, 294.

Hayakawa, T., Windsor, C. R., and Fox, S. W. (1967), *Arch. Biochem. Biophys. 118*, 265.

Hayashi, Y., Teramoto, A., Kawahara, K., and Fujita, H. (1969), *Biopolymers 8*, 403.

Heinrich, M. R., Rohlfing, D. L., and Bugna, E. (1968), *Experientia 24*, 140.

Heinrich, M. R., Rohlfing, D. L., and Bugna, E. (1969), *Arch. Biochem. Biophys. 130*, 441.

Helinski, D. R., and Yanofsky, C. (1966), *Proteins, 2nd Ed., 4*, 1.

Howard, J. C., and Scheraga, H. A. (1972), *Macromolecules 5*, 328.

Ingwall, R. T., Scheraga, H. A., Lotan, N., Berger, A., and Katchalski, E. (1968), *Biopolymers 6*, 331.

Katchatski, E., Sela, M., Silman, H. I., and Berger, A. (1964), *Proteins, 2nd Ed., 2*, 527.

Kauzmann, W. (1959), *Advan. Protein Chem. 14*, 1.

Lakowicz, J. R., and Weber, G. (1973), *Biochemistry 12*,

4171.
Lapanje, S., and Tanford, C. (1967), *J. Amer. Chem. Soc. 89*, 5030.
Laurent, T. C., and Killander, J. (1964), *J. Chromatogr. 14*, 317.
Lewis, P. N., Go, N., Go, M., Kotelchuck, D., and Scheraga, H. A. (1970), *Proc. Nat. Acad. Sci. U. S. 65*, 810.
Lewis, P. N., Momany, F. A., and Scheraga, H. A. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 2293.
Lim, V. I. (1972), *Dokl. Akad. Nauk. SSSR 203*, 224.
Merrifield, R. B. (1969), *Advan. Enzymol. 32*, 221.
Miller, W. G., Brant, D. A., and Flory, P. J. (1967), *J. Mol. Biol. 23*, 67.
Miller, W. G., and Flory, P. J. (1966), *J. Mol. Biol. 15*, 298.
Miller, W. G., and Goebel, C. (1968), *Biochemistry 7*, 3925.
Miller, W. G., and Nylund, R. E. (1965), *J. Amer. Chem. Soc. 87*, 3542.
Morita, K., Simons, E. R., and Blout, E. R. (1967), *Biopolymers 5*, 259.
Nitta, K., and Sugai, S. (1967), *Nippon Kagaku Zasshi 88*, 412.
Ponnamperuma, C. (1972), *in* Exobiology, Ponnamperuma, C., Ed., Amsterdam, North-Holland Publishing Co.
Ptitsyn, O. B., and Finkelstein, A. V. (1970), *Biofizika 15*, 757.
Ptitsyn, O. B., Kron, A. K., and Eizner, Y. Y. (1968), *J. Polym. Sci., Part C 16*, 3509.
Puett, D. (1972), *Biochemistry 11*, 4304.
Quiocho, F. A., and Lipscomb, W. N. (1971), *Advan. Protein Chem. 25*, 1.
Rao, S. P., and Miller, W. G. (1973), *Biopolymers 12*, 835.
Reisner, A. H., and Rowe, J. (1969), *Nature (London) 222*, 558.
Roberts, G. C. K., and Jardetzky, O. (1970), *Advan. Protein Chem. 24*, 447.
Saunders, M. A., and Rohlfing, D. L. (1972), *Science 176*, 172.
Singhal, R. P., and Atassi, M. Z. (1970), *Biochemistry 9*, 4252.
Snipp, R. L. (1964), Ph.D. Thesis, University of Iowa.
Snipp, R. L., Miller, W. G., and Nylund, R. E. (1965), *J. Amer. Chem. Soc. 87*, 3547.
Sugiyama, H., and Noda, H. (1970), *Biopolymers 9*, 459.
Tanford, C. (1961), The Physical Chemistry of Macromolecules, New York, N. Y., Wiley.
Tanford, C. (1962a), *Advan. Protein Chem. 17*, 70.
Tanford, C. (1962b), *J. Amer. Chem. Soc. 84*, 4240.
Tanford, C. (1968), *Advan. Protein Chem. 23*, 121.
Tanford, C. (1970), *Advan. Protein Chem. 24*, 1.
Tanford, C., Kawahara, K., and Lapanje, S. (1966), *J. Biol. Chem. 241*, 1921.
Tanford, C., Kawahara, K., and Lapanje, S. (1967), *J. Amer. Chem. Soc. 89*, 729.
Waugh, D. F. (1954), *Advan. Protein Chem. 9*, 325.
Wetlaufer, D. B. (1962), *Advan. Protein Chem. 17*, 303.
Yang, J. T. (1961), *Advan. Protein Chem. 16*, 323.
Zimm, B. H., and Bragg, J. (1959), *J. Chem. Phys. 31*, 526.
Zimm, B. H., and Rice, S. A. (1960), *Mol. Phys. 3*, 391.

Intensity Fluctuation Spectroscopy of Laser Light Scattered by Solutions of Spherical Viruses: R17, $Q\beta$, BSV, PM2, and T7. I. Light-Scattering Technique†

Peter N. Pusey,* Dennis E. Koppel,‡ and Dale W. Schaefer,§ Rafael D. Camerini-Otero, and Seymour H. Koenig

ABSTRACT: The technique of digital autocorrelation of intensity fluctuations in scattered laser light has only recently been applied to biological problems. It affords precise and rapid measurements of the translational diffusion constants of macromolecules; typically, an accuracy of 1% can be achieved in about 1 min. This paper is intended to be sufficiently detailed and to contain sufficient information to enable workers in the field of biophysical chemistry to assess the applicability of the technique to problems of their interest. We discuss the origin of intensity fluctuations in laser light scattered by a macromolecular solution and give a brief review of the theory of photocount autocorrelation. We then describe the apparatus, and the methods for the reduction of experimental data. In addition to discussing the least-squares data analysis, we present a general method by which various spurious effects such as dust in the sample, and real effects, such as macromolecular polydispersity, may be detected and characterized. Finally we consider sources of error in the measurement of D and present typical data for solutions of R17 virus to illustrate the performance of our apparatus.

The new technique of digital autocorrelation of intensity fluctuations in scattered laser light affords precise and rapid measurement of macromolecular diffusion coefficients (Jakeman and Pike, 1969; Foord *et al.*, 1970; Pusey *et al.*, 1972), values for which can typically be obtained with an accuracy of 1% in about 1 min. We have used this technique to

† From the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598 (P. N. P., D. E. K., D. W. S., and S. H. K.), the Department of Physics, Columbia University, New York, New York 10027 (D. E. K.), and the Department of Molecular Biophysics, The Public Health Research Institute of the City of New York, Inc., New York, New York 10016 (R. D. C.-O.). *Received May 8, 1973.* This research was supported by IBM. R. D. C.-O. is an M.D.–Ph.D. Predoctoral Fellow at the New York University School of Medicine, supported by U. S. Public Health Service Medical Scientist Training Grant No. 5-TO5-6MO-1668-06.

* To whom to address correspondence at: Royal Radar Establishment, Gt. Malvern, Worcs., England.

‡ Present address: Cornell University, Department of Applied Physics, Clark Hall, Ithaca, N. Y. 14850.

§ Present address: Sandia Laboratories, Albuquerque, N. M. 87115.